

Figure 1

SEQ ID NO: 12

GTCATGAAAT TGGAAATCTGA CAAGACGTTT CCAATCATGT TGGGAAGGGAA
GATAAAGGGC TACGCTTGTT TGGTCGAGG GAAGTTATTC AGGCCGATGC
ATGTGGAAGG CAAGATCGAC AACGACGTTT TGGCCGCGCT TAAGACGAAG
AAAGCATCCA AATACGATCT TGAGTATGCA GATGTGCCAC AGAACATGCG
GGCCGATACA TTCAAATACA CCCATGAGAA ACCCCAAGGC TATTACAGCT
GGCATCATGG AGCAGTCCAA TATGAAAATG GGCCTTTCAC GGTGCCGAAA
GGAGTTGGGG CCAAGGGAGA CAGCGGACGA CCCATTCTGG ATAACCAGGG
ACGGGTGTC GCTATTGTGC TGGGAGGTGT GAATGAAGGA TCTAGGACAG
CCCTTTCAGT CGTCATGTGG AACGAGAAGG GAGTTACCGT GAAGTATACT
CCGGAGAACT GCGAGCAATG GTAATGA

SEQ ID NO: 1

VMKLESDKTF PIMLEGKING YACVVGKLF RPMHVEGKID NDVLAALKTK
KASKYDLEYA DVPQNMRA DT FKYTHEKPQG YYSWHHGAVQ YENGRFTVPK
GVGAKGDSGR PILDNQGRVV AIVLGGVNEG SRTALSVVMW NEKGVTVKYT
PENCEQW

Figure 2A

SEQ ID NO: 31

Adaptein-1 nucleotide sequence:

GTCATGAAAT TGGAACTCTGA CAAGACGTTC CCAATCATGT TGGAAAGGGAA
GATAAAGGGC TACGCTTGTG TGGTCGGAGG GAAGTTATTC AGGCCGATGC
ATCTGCAAGC CAAGATCGAC AACGACGTTC TGGCCGCGCT TAAGACGAAG
AAAGCATCCA AATACGATCT TGAGTATGCA GATGTGCCAC AGAACATGCG
GGCCGATACA TTCAAATACA CCCATGAGAA ACCCCAAGGC TATTACAGCT
GGCATCATGG AGCAGTCCAA TATGAAAATG GGCGTTTCAC GGTGCCGAAA
GGAGTTGGGG CCAAGGGAGA CAGCGGACGA CCCATTCTGG ATAACCAGGG
ACGGGTGGTC GCTATTGTGC TGGGAGGTGT GAATGAAGGA TCTAGGACAG
CCCTTTCAGT CGTCATGTGG AACAAGCTTT CTCCACATTA TGCTCAACTC
GAGGGAGTTA CCGTGAAGTA TACTCCGGAG AACTGCGAGC AATGGTAATG
AGC

SEQ ID NO: 32

Adaptein-2 nucleotide sequence:

GTCATGAAAT TGGAACTCTGA CAAGACGTTC CCAATCATGT TGGAAAGGGAA
GATAAAGGGC TACGCTTGTG TGGTCGGAGG GAAGTTATTC AGGCCGATGC
ATCTGGAAGG CAAGATCGAC AACGACGTTC TGGCCGCGCT TAAGACGAAG
AAAGCATCCA AATACGATCT TGAGTATGCA GATGTGCCAC AGAACATGCG
GGCCGATACA TTCAAATACA CCCATGAGAA ACCCCAAGGC TATTACAGCT
CCCATCATCG AGCACTCCAA TATGAAAATG GGCGTTTCAC GGTGCCGAAA
GGAGTTGGGG CCAAGGGAGA CAGCGGACGA CCCATTCTGG ATAACCAGGG
ACGGGTGGTC GCTATTGTGC TGGGAGGTGT GAATGAAGGA TCTAGGACAG
CCCTTTCAGT CGTCATGTGG AACAAGCTTA GAAGCGGTAC TCAATGGCTC
GAGGGAGTTA CCGTGAAGTA TACTCCGGAG AACTGCGAGC AATGGTAATG
AGC

Adaptein-1 protein sequence: SEQ ID NO: 33

VMKLESDKTF PIMLEGKING YACVVGGKLF RPMHVEGKID NDVLAALKTK
KASKYDLEYA DVPQNMRA DT FKYTHEKPQG YYSWHHGA VQ YENGRFTVPK
GVGAKGDSGR PILDNQGRVV AIVLGGVNEG SRTALSVVMW NKLSPHYAQL
EGVTVKYTPE NCEQW

SEQ ID NO: 34

Adaptein-2 protein sequence:

VMKLESDKTF PIMLEGKING YACVVGGKLF RPMHVEGKID NDVLAALKTK
KASKYDLEYA DVPQNMRA DT FKYTHEKPQG YYSWHHGA VQ YENGRFTVPK
GVGAKGDSGR PILDNQGRVV AIVLGGVNEG SRTALSVVMW NKLSRGTQWL
EGVTVKYTPE NCEQW

Alignment of adaptein nucleotide sequences with CCD sequence:

A-1 GTCATGAAAT TGGAAATCTGA CAAGACGTTT CCAATCATGT TGGAAAGGGAA
A-2 GTCATGAAAT TGGAAATCTGA CAAGACGTTT CCAATCATGT TGGAAAGGGAA
CCD GTCATGAAAT TGGAAATCTGA CAAGACGTTT CCAATCATGT TGGAAAGGGAA

A-1 GATAAACGGC TACGCTTGTG TGGTCGGAGG GAAGTTATTC AGGCCGATGC
A-2 GATAAACGGC TACGCTTGTG TGGTCGGAGG GAAGTTATTC AGGCCGATGC
CCD GATAAACGGC TACGCTTGTG TGGTCGGAGG GAAGTTATTC AGGCCGATGC

A-1 ATGTGGAAGG CAAGATCGAC AACGACGTTT TGGCCGCGCT TAAGACGAAG
A-2 ATGTGGAAGG CAAGATCGAC AACGACGTTT TGGCCGCGCT TAAGACGAAG
CCD ATGTGGAAGG CAAGATCGAC AACGACGTTT TGGCCGCGCT TAAGACGAAG

A-1 AAAGCATCCA AATACGATCT TGAGTATGCA GATGTGCCAC AGAACATGCG
A-2 AAAGCATCCA AATACGATCT TGAGTATGCA GATGTGCCAC AGAACATGCG
CCD AAAGCATCCA AATACGATCT TGAGTATGCA GATGTGCCAC AGAACATGCG

A-1 GGCCGATACA TTCAAATACA CCCATGAGAA ACCCCAAGGC TATTACAGCT
A-2 GGCCGATACA TTCAAATACA CCCATGAGAA ACCCCAAGGC TATTACAGCT
CCD GGCCGATACA TTCAAATACA CCCATGAGAA ACCCCAAGGC TATTACAGCT

A-1 GGCATCATGG AGCAGTCCAA TATGAAAATG GGCCTTTTAC GGTGCCGAAA
A-2 GGCATCATGG AGCAGTCCAA TATGAAAATG GGCCTTTTAC GGTGCCGAAA
CCD GGCATCATGG AGCAGTCCAA TATGAAAATG GGCCTTTTAC GGTGCCGAAA

A-1 GGAGTTGGGG CCAAGGGAGA CAGCGGACGA CCCATTCTGG ATAACCAGGG
A-2 GGAGTTGGGG CCAAGGGAGA CAGCGGACGA CCCATTCTGG ATAACCAGGG
CCD GGAGTTGGGG CCAAGGGAGA CAGCGGACGA CCCATTCTGG ATAACCAGGG

A-1 ACGGGTGGTC GCTATTGTGC TGGGAGGTGT GAATGAAGGA TCTAGGACAG
A-2 ACGGGTGGTC GCTATTGTGC TGGGAGGTGT GAATGAAGGA TCTAGGACAG
CCD ACGGGTGGTC GCTATTGTGC TGGGAGGTGT GAATGAAGGA TCTAGGACAG

(HindIII) (XhoI)
A-1 CCTTTTCAGT CGTCATGTGG AAC---AAGCTT TCTCCACATTA TGCTCAA CTCGA
A-2 CCTTTTCAGT CGTCATGTGG AAC---AAGCTT AGAAGCGGTAC TCAATGG CTCGA
CCD CCTTTTCAGT CGTCATGTGG AACGAG-----

A-1 ---GGAGTTA CCGTGAAGTA TACTCCGGAG AACTGCGAGC AATGGTAATGAGC
A-2 ---GGAGTTA CCGTGAAGTA TACTCCGGAG AACTGCGAGC AATGGTAATGAGC
CCD AAGGGAGTTA CCGTGAAGTA TACTCCGGAG AACTGCGAGC AATGGTAATGAGC

Figure 2B

Figure 2C

Alignment of adaptein protein sequences with CCD sequence:

A-1 VMKLESDKTF PIMLEGKING YACVVGGKLF RPMHVEGKID NDVLAALKTK
A-2 VMKLESDKTF PIMLEGKING YACVVGGKLF RPMHVEGKID NDVLAALKTK
CCD VMKLESDKTF PIMLEGKING YACVVGGKLF RPMHVEGKID NDVLAALKTK

A-1 KASKYDLEYA DVPQNMRA DT FKYTHEKPQG YYSWHHGAVQ YENGRFTVPK
A-2 KASKYDLEYA DVPQNMRA DT FKYTHEKPQG YYSWHHGAVQ YENGRFTVPK
CCD KASKYDLEYA DVPQNMRA DT FKYTHEKPQG YYSWHHGAVQ YENGRFTVPK

A-1 GVGAKGDSGR PILDNQGRVV AIVLGGVNEG SRTALSVVMW N-KLSPHYAOLE
A-2 GVGAKGDSGR PILDNQGRVV AIVLGGVNEG SRTALSVVMW N-KLRSGTQWLE
CCD GVGAKGDSGR PILDNQGRVV AIVLGGVNEG SRTALSVVMW NE-----

A-1 -GVTVKYTPE NCEQW
A-2 -GVTVKYTPE NCEQW
CCD KGVTVKYTPE NCEQW

Figure 3

SEQ ID NO: 6

5'ATGTACGGTCGTAAAAACGTCGTCAGCGTCGTCGTGTCATGAAATTGGAA
TCTGACAAGACGTTCCCAATCATGTTGGAAGGGAAGATAAACGGCTACGCTT
GTGTGGTCGGAGGGAAGTTATTCAGGCCGATGCATGTGGAAGGCAAGATCGA
CAACGACGTTCTGGCCGCGCTTAAGACGAAGAAAGCATCCAAATACGATCTT
GAGTATGCAGATGTGCCACAGAACATGCGGGCCGATACATTCAAATACACCC
ATGAGAAACCCCAAGGCTATTACAGCTGGCATCATGGAGCAGTCCAATATGA
AAATGGGCGTTTCACGGTGCCGAAAGGAGTTGGGGCCAAGGGAGACAGCGG
ACGACCCATTCTGGATAACCAGGGACGGGTGGTCGCTATTGTGCTGGGAGGT
GTGAATGAAGGATCTAGGACAGCCCTTTCAGTCGTCATGTGGAACAAGCTTG
GATCTTCTCTCGAGGGAGTTACCGTGAAGTATACTCCGGAGAACTGCGAGCA
ATGGTAA3'.

SEQ ID NO: 7

MYGRKKRRQRRRVMKLESDKTFPIMLEGKINGYACVVGGKLFRPMHVEGKIDN
DVLAAALKTKKASKYDLEYADVPQNMRA DTFKYTHEKPQGYYSWHHGAVQYE
NGRFTVPKGVGAKGDSGRPILDNQGRVVAIVLGGVNEGSRTALSVVMWNEKGV
TVKYTPENCEQW.